

Evaluation Report for the EPA BASINS Decision Support Tool

The BASINS Decision Support Tool has been developed by the U.S. Environmental Protection Agency to simulate watershed pollution and to identify sources of nonpoint pollution.

Hydrological Sciences Branch
NASA Goddard Space Flight Center (Code 614.3)
Greenbelt, MD 20771

Acknowledgments

This work was supported by the NASA Applied Science Program. The work has been conducted at both NASA Goddard Space Flight Center and at partnering universities and the Chesapeake Bay Program.

EPA BASINS Teams

David Toll–*Team Leader*

NASA GSFC

David Toll

Edwin Engman

Kristi Arsenault

Joseph Nigro

Jonathan Triggs

Hunter College

Wenge Ni-Meister

University of Maryland Center for
Environmental Studies (UMCES)

Keith Eshleman

U.S. EPA

Ed Partington

Jim Carleton

Amy Neuman

Chesapeake Bay Program

Gary Shenk (USGS)

Angelica L. Gutiérrez-Magness
(UMCP/USGS).

University of Wisconsin

Phil Townsend

TABLE OF CONTENTS

Table of Contents	3
Tables	4
Figures.....	4
Executive Summary	5
1.0 Introduction.....	6
1.1. NASA Mission Traceability	6
1.2. Nonpoint Source Pollution	7
1.3. The BASINS DST.....	7
1.4. Systems Engineering Approach	9
2.0 Selection of the EPA Decision Support Tool	10
2.1. Overview of EPA BASINS	10
2.2. Selection of Appropriate DST	11
2.3. Approach to Improving HSPF Results	12
3.0 Consideration of NASA Inputs.....	13
3.1. Overview	13
3.2. NASA Satellite Derived Products	14
3.3. NASA Model Derived Products	14
3.4. Potential for NASA Data for use in HSPF	15
3.4.1. Improved Model Parameters.....	15
3.4.2. Improved Model Forcing	16
3.4.3. Improved Initial Conditions.....	17
4. Gaps in Meeting BASINS Needs.....	19
5.0 Partnering.....	20
5.1 Overview	20
5.2 Developing Partners	20
6.0 Conclusions and Recommendations	22
6.1. Findings.....	22
6.2 Recommendations.....	22
6.3. Next Steps	23
7.0 References.....	24
7.0 Abbreviations and Acronyms	26
Appendix A. The Land Information System	28

TABLES

Table 1. MODIS TERRA and AQUA data product list for evaluation.....	18
Table 2. Typical LIS forcing and output products	19

FIGURES

Figure 1. BASINS operational Overview	8
Figure 2. The illustration of the systems engineering architecture underlying the Applied Science Program and its specific application to the EPA BASINS Decision Support Tool.	9
Figure 3. Summary of NASA data sets for input to BASINS-HSPF.	17

EXECUTIVE SUMMARY

The Environmental Protection Agency (EPA) estimates that over 20,000 bodies of water throughout the country do not meet water quality standards. In recent U.S. opinion polls, water quality is in the top 3 of 5 environmental concerns. Nonpoint sources -- pollution from urban and agricultural land that is transported by runoff -- typically cause 90 percent of impairments. Ideally, water quality monitoring should occur at numerous locations within a watershed on a continuous basis to assess fluctuations in water quality under different flow and seasonal conditions in order to identify pollution sources. Unfortunately, states lack the significant resources needed to assess and manage water bodies by monitoring alone. To overcome these shortcomings, the EPA has developed the BASINS (Better Assessment Science Integrating Point and Nonpoint Sources) modeling system for performing watershed and water quality studies. BASINS is a decision support tool used to assess point source and non-point source pollution by states, tribes, and regional and local agencies use for making watershed and water quality assessment studies. The key to this suite of models is the Hydrological Simulation Program - Fortran (HSPF), which calculates diurnal stream flow rates and the corresponding pollutant concentrations at the watershed outlet. HSPF does not perform well when quality spatial data are not readily available.

In response to an urgent and strong need for more spatially- and temporally- complete information, EPA has partnered with NASA to use their high spatial and temporal hydrological variables (e.g., precipitation, evaporation, etc.) with initial conditions (e.g., snow cover) from the NASA Land Information System (LIS) and land use/land cover derived from a suite of satellite data. The collaborative work is to conduct systems engineering related analyses, including “Evaluation”, “Verification and Validation”, and “Benchmarking” toward an operational approach of implementing NASA data in the current EPA water quality assessment decision tool. **The purpose of this report is to provide an “Evaluation” of the NASA satellite and modeling products that may improve the BASINS DST.** We are evaluating primarily the following NASA products:

- **LIS Precipitation:** derived from gauge data, NOAA radar and NASA-NOAA satellite data, to provide high spatial and temporal data to drive HSPF.
- **LIS Evapotranspiration and Snow Water Equivalence:** calculated based on complex physical model, are fundamental water availability variables of HSPF and can be used to better initialize HSPF soil water storage conditions
- **Satellite Land Use and Land Cover:** derived from MODIS, ASTER and ALI data, to provide phenological and other analyses critical to monitor water quality under different environmental/flow conditions and to help identify pollution sources.

The work will emphasize the infusion of NASA data in to the “BASINS-HSPF” decision support tool and an improved version of HSPF used by the Chesapeake Bay Program (CBP). The unique capabilities provided by NASA satellite remote sensing and modeling have significant potential to provide better input data and initial conditions to assist the EPA in meeting water quality goals. The project attempts to leverage the large investment in the NASA Earth-Sun Science data to a federal agency with national applications that may provide a significant return for policy making affecting peoples’ every day lives. The benefit to EPA would be the improvement of watershed assessment through the adaptation and inclusion of state-of-the-art Earth systems data while NASA performs its mission “to understand and protect our home planet.”

1.0 INTRODUCTION

1.1 NASA Mission Traceability

The NASA vision and mission statements include a clear focus on the Earth and life on Earth. NASA seeks to improve life on Earth by enabling people to use measurements of our home planet in valuable ways to manage our natural resources. NASA's Earth Science Division has primary responsibility for two Agency-wide, Earth oriented themes in the NASA strategic plan: Earth system science and Earth science applications. In serving these themes, the division works with its domestic and international partners to provide accurate, objective scientific data and analysis to advance our understanding of Earth system processes and to help policy makers and citizens achieve economic growth and effective, responsible stewardship of Earth's resources.

The Earth Science Applications Program has as its primary goal to extend the benefits of NASA's Earth science to the broader community. To do this, NASA has identified twelve applications of national priority of which water management is one. The Water Management Program Element extends products derived from Earth science information, models, technology and other capabilities into partners' decision support tools to help them meet their water management responsibilities and mandates to support water resource managers. The general areas related to water availability and quality includes the following.

- Estimating water storage – snowpack, soil moisture, aquifer volumes
- Modeling and predicting water fluxes - Evapotranspiration, rain, runoff
- Water quality – turbidity, temperature, modeling nonpoint source pollution

It is in response to this last item, nonpoint source pollution, that NASA is partnering with the Environmental Protection Agency (EPA) to investigate the feasibility of using NASA data to improve EPA's capability to model watershed nonpoint source pollution. The EPA is responsible for protecting various bodies of water in the U.S. The primary guideline for EPA's mandate is the Clean Water Act of 1972. One of the regulations spelled out in this Act is that EPA must track the Total Maximum Daily Load (TMDL) for any watershed. The TMDL defines the amount of pollution that can be carried by water before it is determined to be "polluted". There are essentially two ways for EPA to monitor this variable: one is through in-stream measurements and sampling, and two, through modeling the streams response to storm runoff and pollution loadings. The first option would be prohibitively expensive and impractical for the entire U.S.. The modeling approach is the only practical solution. To do this, EPA developed the BASINS decision support tool.

The problem of nonpoint source pollution is a spatially and temporally complex issue. Currently, pollution monitoring is performed at a very limited number of ground stations in the U.S. Many important watersheds have no monitoring of pollution transport or streamflow at all. The models in BASINS currently rely on point-based meteorological and pollution measurements. By incorporating NASA remote-sensing data, many of the critical input variables to BASINS can be improved spatially. Satellite gridded data and data products will enhance BASINS output results, thereby leading to better decisions regarding water quality, and therefore improved management of the nation's water resources. This

goal complements the NASA Mission Statement “To understand and protect our home planet...” and NASA’s Vision “to improve life here...”

1.2. Nonpoint Source Pollution

The Environmental Protection Agency (EPA) estimates that over 20,000 bodies of water throughout the country are too polluted to meet water quality standards (USGAO, 2000). Included in this figure are more than 300,000 mile of river and shorelines and 5 million acres of lakes. The Clean Water Act requires states to identify water bodies that do not meet applicable standards and establish a pollution budget known as the Total Maximum Daily Load (TMDL) for each pollutant contributing to the degraded waters. A TMDL is a sum of the maximum allowable loads of a pollutant from point and nonpoint sources that a water body can receive and still meet its water quality standards. Typical constituents contributing to water quality impairment are: pathogens, heavy metals (e.g., mercury, copper, lead), nutrients (e.g., nitrogen and phosphorus), toxic organics, enriched biochemical oxygen demand, and low concentrations of dissolved oxygen. Nonpoint sources, primarily pollution from urban and agricultural land that is transported by precipitation and runoff either as a sole source or in conjunction with point sources, cause 90 percent of EPA section 303d listed impairments.

Nonpoint source pollution from agriculture, logging, and urbanization is the leading cause of degraded water quality in the U.S. (Smith et al., 1987). The retention and transport of nonpoint source pollutants depend upon multiple features of a landscape, including land cover, soils and underlying geology, surface topography, and stream network characteristics. As a consequence, the effective modeling of water quality effects of nonpoint sources requires adequate spatial characterization of soils, topography and land cover.

1.3. The BASINS DST

Ideally, one would like to monitor water quality at numerous locations within a watershed on a periodic basis to assess fluctuations in water quality under different flow and seasonal conditions and assist in the identification of pollution sources. Unfortunately, states lack the resources to assess and protect water bodies with monitoring data alone. To overcome this shortcoming, the EPA has developed a modeling system for performing watershed and water quality studies. BASINS (Better Assessment Science Integrating point and Nonpoint Sources) (USEPA, 2001) is a multipurpose environmental analysis system to assist regional, state and local agencies in their assessment obligations. BASINS is designed to evaluate environmental and ecological conditions in a watershed context. BASINS is also configured to develop TMDLs for water bodies that are not meeting water quality standards. Developing TMDLs requires a watershed-based approach that integrates both point and nonpoint sources.

BASINS includes a suite of models designed to model meteorological conditions, flow across watersheds, and ultimately pollutant transport. The systems overview for BASINS is illustrated in Figure 1. The results produced by the models enable more accurate understanding of conditions leading to excessive TMDL values. BASINS includes

hydrologic and pollutant fate and transport models that simulate streamflow and runoff from the land surface (nonpoint sources). Accuracy in modeling streamflow and runoff is essential for estimating water quality and establishing TMDLs at locations within a watershed. Quantitative measures or estimates of streamflow are needed to define concentrations of water quality constituents.

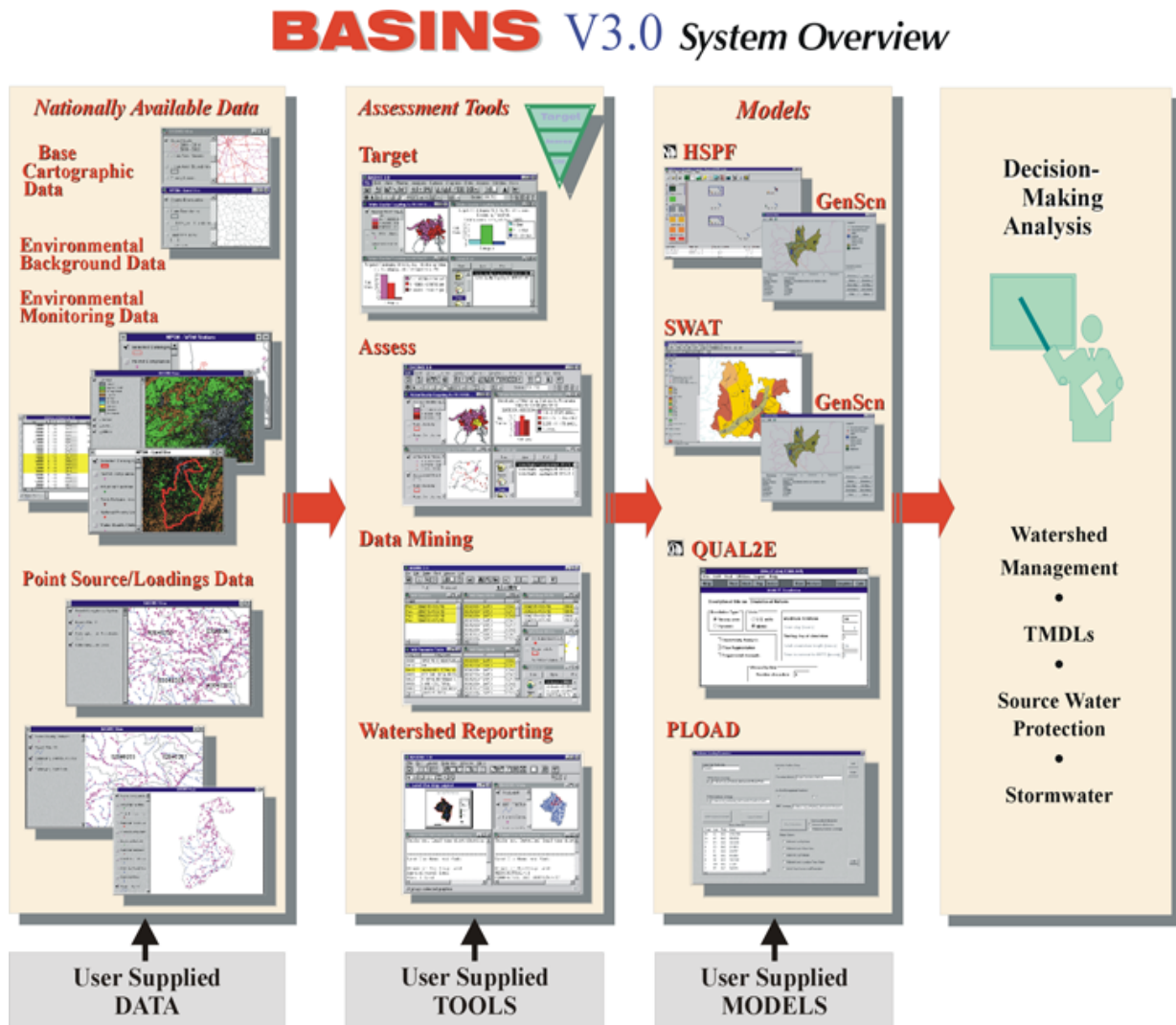


Figure 1. BASINS operational overview.

1.4. Systems Engineering Approach

The Earth Science Applications program's approach to extend the benefits of Earth science observations and predictions to decision-support tools is based on fundamental system engineering principles. Figure 2 illustrates the architecture underlying the activities of the Earth Science Applications program. To the right, partner agencies own, develop and operate decision support tools to carry out their water management mandates. On the left, NASA extends the observations, model predictions, and computational techniques from its Earth science research to support its partners.

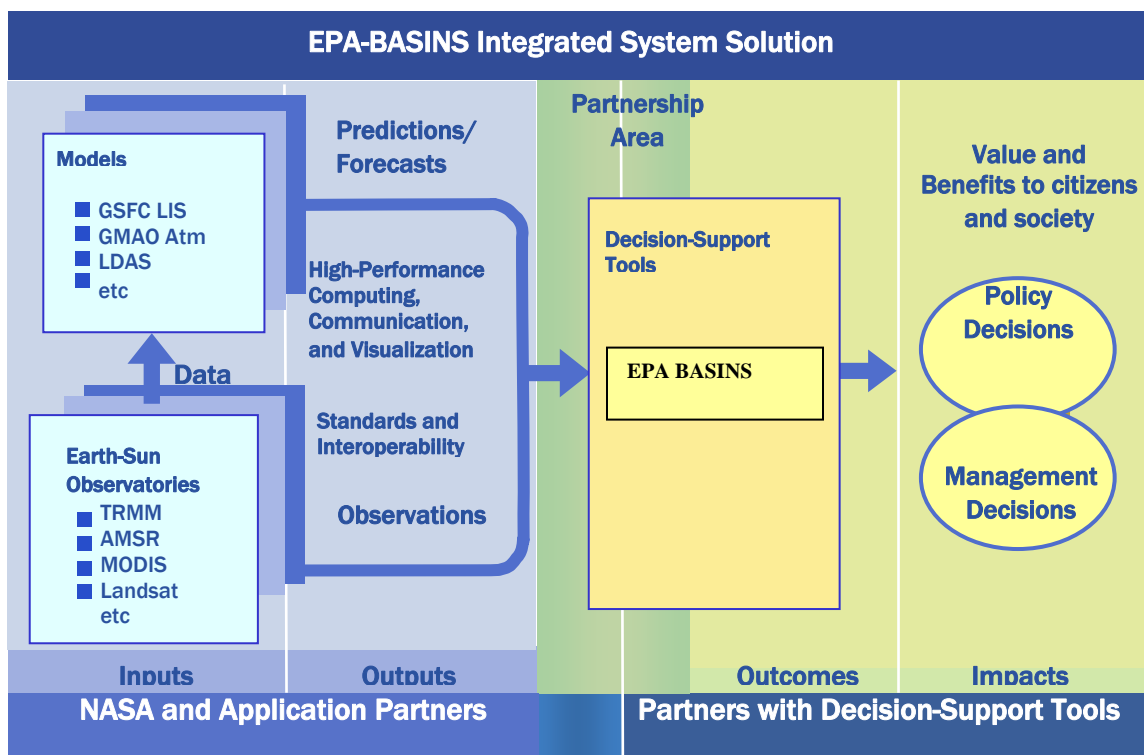


Figure 2. Illustration of the systems engineering architecture underlying the activities of the Earth Science Applications Program and its specific application to the EPA BASINS Decision Support Tool.

The systems engineering approach involves the four steps of *evaluation*, *validation*, *verification and benchmarking* to test the utility of NASA Earth science data for improving the performance of EPA's watershed and water quality decision support tools. The emphasis of this report is to *evaluate* the use of NASA data products through study of the EPA BASINS Decision Support Tool (DST). The benefit to EPA would be improvement of watershed assessment through the adaptation and inclusion of state-of-the-art Earth systems

data and data products while a benefit for NASA benefit would stem from continuing its mission “to understand and protect our home planet.”

The major emphasis is the use of NASA products to estimate important model parameters (e.g., land use, buffer zones, etc.), improve forcing functions (e.g., precipitation, evaporation, etc.) and provide initial conditions (e.g., snow cover, etc.) to improve the performance and accuracy of BASINS. The *evaluation* step in this process was to assess BASINS inputs and outputs. Next, NASA modeling and remote sensing products were matched against the existing inputs to BASINS.

Following this initial evaluation, the most promising NASA data products will be substituted into BASINS one at a time to test for improvements in HSPF-simulated stream flow. The process of ingesting NASA data into BASINS constitutes the second step of the systems engineering approach. This *Verification and Validation* phase involves the development of techniques for using NASA data in BASINS. Also, the *baseline* data will be defined and *benchmark* metrics will be developed within this phase.

The third and final phase in the systems engineering process, the *benchmarking* phase, will include the results of testing each NASA input separately against the established baseline in phase 2. Also, the *benchmarking* phase will describe the processes necessary to integrate results of this effort into everyday BASINS use at the EPA and partner level.

An anticipated outcome of this project is the determination of optimal data sets for use with watershed assessment tools. A key part of the *benchmarking* procedure will be comparisons of results using EPA traditional data and configurations versus those with NASA data and to document the improvements with quantitative measures against the baseline results.

2.0 SELECTION OF THE EPA DECISION SUPPORT TOOL

2.1 Overview of EPA BASINS

BASINS (Better Assessment Science Integrating Point and Nonpoint Sources) is a multipurpose environmental analysis decision support tool (DST) used to assist regional, state and local agencies in their water quality assessment obligations. BASINS was developed by the EPA to meet three objectives:

1. To facilitate examination of environmental information
2. To support analysis of environmental systems
3. To provide a framework for examining management alternatives.

BASINS consists of several components that include data base elements, data mining tools, watershed and water quality models, and reporting tools. These component parts are integrated into the DST using a Geographical Information System (GIS) program, Arc View.

BASINS is configured to support environmental and ecological studies in a watershed context. BASINS is also configured to develop Total Maximum Daily Loads (TMDLs) for

water bodies that are not meeting water quality standards. Section 303(d) of the Clean Water Act requires state to develop TMDLs for water bodies that are not meeting applicable water quality standards. Developing TMDLs requires a watershed-based approach that integrates both point and nonpoint sources.

The lack of widespread use and acceptance of BASINS can be attributed to a lack of good meteorological, hydrologic and water quality data to monitor a watershed. EPA has sponsored the development of a continuous hydrologic simulation model known as HSPF (Hydrologic Simulation Program - FORTRAN) (Donigian et al., 1995 and Bicknell et al., 1997). HSPF simulates nonpoint source runoff and pollution loadings for a watershed, combines these with point source contributions, and performs flow and water quality routing in the watershed channels.

2.2 Selection of Appropriate DST

The BASINS system combines six components to provide the range of tools needed for performing watershed and water quality analysis. These interrelated components can be summarized as follows:

1. National environmental data bases (basic cartographic data, environmental background and monitoring data, point sources/loading data)
2. Assessment tools (TARGET {broad based, preliminary conclusions}, ASSESS {status of specific stream reaches and evaluate the need for source characterization and cause-effect relationships} and Data Mining)
3. Utilities (a series of tools for managing data, delineating sub-watersheds, reclassification of data and overlaying data)
4. Watershed characterization reports (point sources, land use, topography, etc.)
5. Water quality stream models (QUAL2E)
6. Watershed models (HSPF, SWAT, PLOAD)

The decision was made to evaluate the EPA models (5&6 above) to select the optimal opportunities for infusing NASA data and data products with the hope of improving the usefulness and performance of the EPA BASINS system.

QUAL2E is a one dimensional model that analyzes the fate and transport of pollutants selected stream reaches. QUAL2E is best used where you are concerned *with a Dissolved Oxygen (DO) endpoint in an effluent dominated system and can accept the steady state assumptions*. The details and scale of this model eliminated it from further consideration for NASA contributions. Our focus then concentrated on the three watershed models (#6 above).

In considering what strengths a potential NASA contribution could make to improving the application of BASINS to different physiographic regions, we focused on the spatial and temporal characteristics of remote sensing data and data products. EPA considered a continuous simulation model to be critical for a realistic representation of watershed processes. A continuous simulation model automatically takes into consideration the serial

correlation present in flows and other variables, as well as the cross-correlations between measured variables. Based on this criterion, we eliminated PLOAD and SWAT from consideration. PLOAD is a simple watershed model that is based on annual precipitation, land use and Best Management Practices (BMP). PLOAD can be used when you want *estimates of annual and seasonal loading to drive simple eutrophication models*. SWAT is a daily time step model that can predict the effects of land use management and can be used where there are *no nearby meteorological stations with hourly data and where there is no nearby gauged watershed*.

The process of elimination and the matching of NASA capabilities and BASINS needs has led us to focus on the Hydrologic Simulation Program – FORTTRAN (HSPF) (Donigian et al., 1995 and Bicknell et al., 1997) model. HSPF simulates the hydrology and associated water quality processes on pervious and impervious land surfaces and in streams and well mixed impoundments. HSPF is a lumped parameter, continuous streamflow simulation model based on the Stanford Watershed Model (SWM), the first complete watershed model performed on a digital computer. The model requires land use, channel reach, and meteorological data and information on expected pollutants. HSPF is designed to interact with BASINS utilities and data sets to facilitate the extraction of appropriate information and the preparation of model input files. HSPF can be run on a single watershed or a system of multiple connected sub-watersheds that have been delineated using the BASINS “Watershed Delineation” tool and GIS elevation datasets such as the Digital Elevation Model (DEM) provided by USGS. Generally, spatial variability within a large watershed is dealt with by subdividing the watershed into sub-watersheds. In doing this one then must select parameters for each sub-watershed to reflect the spatial heterogeneity.

2.3 Approaches to Improving HSPF Results

In choosing to work with HSPF, we realized that the NASA impact could be derived from all three potential areas in which NASA data and science products may be used to improve the BASINS model performance. These include:

- Improved parameters (i.e., land use, buffer zones, from satellite imagery, etc.)
- Improved forcing (i.e., spatially distributed precipitation, evaporation, wind, solar radiation, etc. derived from data assimilation)
- Improved initial conditions (i.e., snow cover, soil moisture, from data assimilation products, etc.)

Improved parameters can take the form of GIS datasets currently available to BASINS and are from many sources, including the National Elevation Data set (Gesch et al., 2002), the National Land Cover Data set (Vogelman et al., 2001), and the STATSGO soils database (USDA, 1993). However, numerous alternative data sets exist, including digital elevation data from the Shuttle Radar Topography Mission (SRTM, e.g., Smith and Sandwell, 2003), soil properties maps (e.g., Hargrove and Luxmoore, 1998), ecoregion delineations (Hargrove and Hoffman, 2004), as well as detailed land use and land cover maps with more hydrologically meaningful categories, such as impervious surface area (e.g., Civco et al., 2002; Wang and Zhang, 2004; and Jantz et al., 2004) or MODIS-derived measures of

vegetation cover and phenology (Hansen et al., 2002; Zhang et al., 2003; and Ni-Meister and Tomita, 2005).

Many of these data sets – especially those related to land cover – are expected to provide more accurate representations of the surface properties within watersheds. Specifically, the dynamic characterization of land cover through time will be an improvement over static classifications. Likewise, the assessment of total imperviousness within a watershed (where every pixel exhibits a range of imperviousness) will be more useful than the simple quantification of pixel area mapped as an impervious class (e.g., “urban or built-up” in the Anderson Level II scheme (Anderson et al., 1976).

Improved forcings for HSPF will concentrate on improving the accuracy of meteorological data at appropriate temporal and spatial resolutions to ensure the quality of the modeling results. Typically hourly station data maintained by NOAA or other organizations is used in HSPF modeling. However, there are many instances in which there are no nearby meteorological data available from ground-based stations for a watershed of interest. In such instances, estimates are usually made by using data from the closest stations. Meteorological data plays a crucial role in simulating stream flow and runoff, which in turn have a significant impact in estimating total pollutant loads and developing TMDLs. Having accurate hourly meteorological data on a relatively small spatial scale could improve HSPF modeling efforts by decreasing modeling uncertainty, increasing the accuracy of TMDL estimates, and allowing for modeling on smaller, sub-watershed scales. More local scale modeling could lead to more efficient placement of Best Management Practices (BMPs) used to control nonpoint source pollution, thereby providing better water quality results at lower costs.

Improved initial conditions involve quantifying the hydrologic status of the watershed at the beginning of the simulation run. Typically these would include variables such as soil moisture, snowpack volume and water content and impoundment levels. Soil moisture is a product derived from data assimilation and in the future from direct satellite measurements. However, it cannot be used to improve HSPF because the soil moisture related parameters in HSPF are simply parameters and are not based on actual levels of soil moisture. However, snow products from data assimilation and satellites have the potential for significant improvements in simulating runoff from snowmelt or rain on snow events.

3.0 CONSIDERATION OF NASA INPUTS

3.1. Overview

There are several potential areas in which NASA data and science products may be used to improve BASINS-HSPF model performance. These include improved parameters, improved forcing, and improved initial conditions. NASA data will be evaluated within the context of a systems engineering approach where NASA data products will be verified and validated, and improvements to BASINS-HSPF will be benchmarked. We summarize next the various NASA data products that may be useful for improving estimated stream flow in BASINS HSPF.

3.2. NASA Satellite Derived Data

The primary NASA satellite data for evaluation is from MODIS. MODIS is the flagship sensor on Terra and Aqua satellite providing a range of land, atmosphere and water products. Table 1 summarizes the MODIS data we plan to study.

Other non-MODIS, but NASA products may be evaluated over the course of the project. For example:

- 1) EO-1 Advance Land Imager (ALI) provides near-Landsat bands with three additional bands. However the ALI does not have a thermal band. The EO-1 satellite follows the Landsat overpass by 1-minute. The EO-1 has a 37 km swath in comparison with the 185 km Landsat ETM+ swath. EO-1 can be tilted for off-nadir scans. EO-1 was launched on November 21, 2000;
- 2) EO-1 Hyperion provides 220 bands between 0.4-2.4. The swath is within an ALI swath, but is only 7.7 km wide. Hyperion can do more biogeochemistry than Landsat TM, including applications such as mine tailings, water quality, plant stress, and species differentiation;
- 3) ASTER has 4 (VNIR) 15 m bands, 6 (SWIR) 30 m bands and 5 (TIR) 90 m bands. ASTER includes multispectral thermal measurements between 8-12 μm . The swath is 60 km. It is similar to Landsat, but can do backward viewing for stereoscopic observation. ASTER provides many of the capabilities of Landsat, but data availability is more limited. ASTER was launched with MODIS on EOS Terra on December 18, 1999;
- 4) The Shuttle Radar Topography Mission (SRTM) has been completed, and has surface elevation data available for the U.S. at 30 m with vertical resolution accuracy of approximately 6-7 m;
- 5) AMSR-E (Advanced Microwave Scanning Radiometer for EOS (AMSR-E) is a 12 channel, six-frequency passive microwave radiometer system that may be used to provide estimates of soil moisture and snow water equivalent.

3.3. NASA Model Derived Products

The primary NASA data sources for evaluation of HSPF parameters, forcing, and initial conditions are NASA's Land Information System (LIS) and the Moderate Resolution Imaging Spectroradiometer (MODIS). MODIS is the flagship sensor on the Terra and Aqua satellite providing a range of land, atmosphere and water products. Table 1 summarizes the MODIS data that will be tested.

NASA's Land Information System (LIS) modeling capabilities, in combination with NASA satellite data products, such as MODIS land surface temperature, are directed toward capturing the most realistic representations of land surface dynamics and their interactions

with the atmosphere over large areas and at high resolutions. The main hydrometeorological variables that LIS produces include soil moisture, evaporation, snow cover, runoff, precipitation, and also radiation and energy budget variables, and most of all of these in some capacity can be used in the BASINS HSPF. Many of the relevant LIS modeled output variables are listed and summarized in Table 2 below.

The NASA North America Land Data System (NLDAS; Mitchell et al., 2004, Cosgrove et al., 2003) forcing data fields, which drive the LIS LSMs, can be used for real-time and retrospective simulations for within HSPF. Also, different atmospheric forecast fields can be used for HSPF and are currently available from LIS and the NLDAS forcing fields (i.e., NCEP Eta-12 km, FSL RUC-20 km). All three modes (i.e., retrospective, realtime and forecast) can be used to generate soil moisture, evapotranspiration and heat fluxes out to 72 hours. Also at NASA, diurnal (2-5 days) to intraseasonal (30-90 days) predictions of precipitation and temperature, generated by the NASA Global Modeling and Assimilation Office (GMAO), could be tested and validated in the DSTs.

3.4. Potential for NASA Data for use in HSPF

As stated above, we anticipate that the NASA impact could be derived from all three potential areas in which NASA data and science products may be used to improve the BASINS model performance. These include 1) Improved Parameters; 2) Improved Forcing; and 3) Improved Initial Conditions:

3.4.1 Improved Model Parameters

Model parameters are considered to be static input values required for running the HSPF model. Variables such as elevation, land use, land cover, and soil properties generally do not change rapidly, but have a great impact on model simulations. However, it is generally believed that using static land cover/land use data can lead to erroneous model simulations. With NASA data, it is possible to update these data more frequently and to even look at seasonal effects as well as episodic events such as fires and insect infestations. The following parameters may be tested using NASA products in place of existing BASINS-HSPF datasets to assess possible improvements in stream flow simulation:

Land Use/Cover and Imperviousness - The high temporal resolution of MODIS (daily at 250m) and high spatial resolution of Landsat (16 days at 15m-60m) enables us to track long and short term trends in land use/land cover. We are also able to look at vegetative continuous fields and phenological information (such as LAI during growing season and fallow cycles) from both MODIS and Landsat, such as LAI during growing and fallow cycles.

Digital Elevation Model – The Shuttle Radar Topography Mission (SRTM) provides a 30m spatial resolution digital elevation model for the U.S. with a vertical resolution accuracy of approximately 6-7m.

3.4.2. Improved Model Forcing

In addition to improved parameters, many of the HSPF model forcing datasets may be substituted with improved NASA data. The HSPF model requires meteorological input data to drive stream flow simulations. All current forcing data within BASINS and HSPF is provided by weather stations, which tend to be widely spaced. The greatest potential improvement in model forcing datasets could be through inclusion of NASA's gridded meteorological datasets such as those provided by LIS. The following variables will be tested in BASINS-HSPF:

Precipitation - A primary forcing in watershed flow estimation is precipitation. NASA has gridded datasets such as from the Tropical Rainfall Measurement Mission (TRMM) and LIS that can provide more accurate spatial representations of precipitation than the current rain-gauge datasets. LIS provides a gridded precipitation product at 1/8th degree resolution, and can provide scaled precipitation to even finer spatial resolutions.

In addition to LIS output, the NOAA Stage 2 NEXRAD Doppler radar product, used as a forcing variable for LIS, can provide gridded hourly precipitation accumulations at a scale of approximately 4km. The primary advantage of using gridded data is to fill in data gaps between rain gauge stations, which are often outside of the test watershed. Because rainfall is a primary driver of the water cycle, accurate representation of rainfall in the HSPF model is critical.

In the future, NASA's Global Precipitation Measurement (GPM) Mission will provide gridded precipitation measurement similar to TRMM. Unlike TRMM, however, GPM will have global coverage, allowing areas north and south of 40 degrees latitude spatially accurate precipitation measurements. The methods of incorporating Stage 2 and LIS precipitation into BASINS-HSPF will allow future satellite-based precipitation datasets to be used by EPA and other partners to improve model forcing.

Evapotranspiration - Many other forcing parameters required to run HSPF may be improved by using NASA datasets in place of existing methods. For example, LIS provides data such as evapotranspiration, air temperature, dew point temperature, wind, and solar radiation. LIS obtains these variables from a variety of sources including satellites and ground-based sensors. LIS provides these datasets at spatial scales down to 1km, and could even produce data at finer resolution in the future. The amount of cloud coverage can affect energy input into HSPF, which can affect other factors such as evapotranspiration and soil temperature. Cloud cover forcing data can be obtained via MODIS and LIS. Preliminary results using LIS derived ET in place of the overly simple models in HSPF lead to improved streamflow simulations.

Air temperature - HSPF uses a simple degree day approach to simulate snow melt. LIS air temperature products have the potential to improve the snow melt simulations, especially in watersheds with no local temperature data.

3.4.3. Improved Initial Conditions

There is an excellent opportunity for NASA data and data products to have an impact on HSPF through improved definition of snow extent and snow water equivalent. Although HSPF does have a simple snow melt algorithm (a degree day approach) it is generally believed that the weakest aspect of the snow simulations come from a lack of knowledge of initial conditions. Satellite and LIS snow extent and snow water equivalent products should go a long way to improving HSPF simulations under these conditions.

Figure 3 below summarizes the opportunities for NASA data and data products to have a positive impact of HSPF simulations.

		FORCING VARIABLES							PARAMETERS			INITIAL CONDITIONS		
		Precipitation	Pot. ET	Air Temperature	Wind Speed	Solar Radiation	Dewpoint Temp.	Cloud Cover	Land Use/Cover	Elevation	Soil Properties	Snow	Water Balance	Soil Temperature
												Snow Accumulation, Snow Melt	Soil Moisture, Soil Moisture Change, ET	
NASA DATASETS	SATELLITE DATA													
	MODIS							●	●			●		
	Landsat TM								●					
	ASTER								●	●				
	SRTM									●				
	TRMM	●												
	MODEL DATA													
LIS	●	●	●	●	●	●	●	●		●		●	●	●

Figure 3. Summary of NASA data sets for input to BASINS-HSPF.

Table 1. MODIS TERRA and AQUA data product list for evaluation.

Land	Description
MOD 12Q1– Land Cover Type	Spatial Resolution: 1 km; Temporal Resolution: 96-day Land Cover Classification Product consisting of 17 classes of land cover representative of the IGBP global vegetation classification scheme and including classes for natural vegetation, developed land, permanent snow and ice, barren and sparsely vegetated land, and water. Additional data layers include the UMD modification of the IGBP Classification scheme, the MODIS LAI/FPAR scheme, and the MODIS Net Primary Production scheme.
MOD 15A2 – Leaf Area Index & FPAR	Spatial Resolution: 250 m; Temporal Resolution: 16-day The Vegetation Indices are transformations of the red, near-infrared, and blue bands to indicate the amount of vegetation present on the ground and to allow for the analysis of spatial and temporal variations in vegetation.
MOD 44A – Vegetation Cover Conversion	Spatial Resolution: 1 km; Temporal Resolution: 8-day The Leaf Area Index defines the structural property of a plant canopy while the Fraction of Photosynthetically Active Radiation measures the proportion of available radiation absorbed by the canopy. Both products have been used for the calculation of surface photosynthesis, evapotranspiration, and annual net primary production.
MOD 13Q1 – Gridded Vegetation Indices (Max NDVI & Integrated MVI)	Spatial Resolution: 250 m; Temporal Resolution: 96-day The vegetation cover conversion product shows the global distribution of where vegetation cover change is occurring at 3-month intervals. An interannual product is also available for analysis of global vegetation-cover change between years.
MOD 10 – Snow Cover	Spatial Resolution: 500 m; Temporal Resolution: 96-day Global snow cover product mapped daily but distributed as 8-day composites.
MOD 11A1 – Land Surface Temperature/Emissivity	Spatial Resolution: 1 km; Temporal Resolution: Daily Land surface temperature product, extracted in Kelvin consisting of daily daytime and nighttime temperatures. The emissivity component reveals the ability of a surface to emit heat by radiation. This product uses algorithms based on geolocation, radiance, cloud masking, atmospheric temperature, water vapor, snow, and land cover information, derived from other MODIS products. ET and snow and ice melt are affected by fluctuations in surface temperature.
MOD 43B3 – Albedo	Spatial Resolution: 1 km; Temporal Resolution: 16-day Albedo quantifies the incident radiation that is reflected by a surface and is dependent on the Bidirectional Reflectance Distribution Function (BRDF) since it is related to land surface reflectance by directional integration. The spatial and temporal distribution of land surface structure and optical properties, which determine albedo, may be due to meteorological parameters, such as soil wetness and snowfall distribution.

Table 2. Typical LIS forcing and output products.

ATMOSPHERIC		LAND SURFACE AND SUBSURFACE	
Net Shortwave Radiation (W/m^2)		Snowpack (kg/m ²)	Water Equivalent
Net Longwave Radiation (W/m^2)			Top 1 m Soil Moisture (kg/m ²)
Downward Solar Radiation Flux (W/m^2)		Snow Depth (m)	Layer 2 Soil Moisture (kg/m ²)
Downward Longwave Radiation Flux (W/m^2)		Snow Cover (%)	Layer 3 Soil Moisture (kg/m ²)
Snowfall, Frozen Precipitation (kg/m ²)		Snowmelt (kg/m ²)	Total Soil Column Wetness (%)
Rainfall, Unfrozen (kg/m ²)		Surface Runoff (kg/m ²)	Root Zone Wetness (%)
Surface Pressure (Pa)		Subsurface Runoff (kg/m ²)	Root Zone Soil Moisture (kg/m ²)
Air Temperature, 2m (K)		Average Sfc Temperature (K)	Total Column Soil Moisture (kg/m ²)
Specific Humidity, 2m (kg/kg)		Surface Albedo (%)	Plant Canopy Surface Water Storage (kg/m ²)
U Wind Component (m/s)		Canopy Surface Water	Canopy Transpiration (W/m^2)
V Wind Component (m/s)		Vegetation Greenness (%)	Aerodynamic Conductance (m/s)
Convective Precipitation (kg/m ²)		Leaf Area Index	Canopy conductance (m/s)
		Evaporation (W/m^2)	Sensible Heat Flux (W/m^2)
		Deep Soil Temperature (K)	Latent Heat Flux (W/m^2)
		Canopy Temperature (K)	Ground Heat Flux (W/m^2)

4.0 GAPS IN MEETING BASINS NEEDS

BASINS is a multipurpose GIS-based interface to water quality modeling. The current reliance on ground station-based input data leaves a great deal of uncertainty in modeled results. In addition, the fact that HSPF is a lumped parameter model in which its many (20+) parameters have no physical meaning and are merely fitting parameters, limits the potential for using physical measurements of hydrological variables to improve its performance.

While NASA remote-sensing and modeling data show promise in improving spatial accuracy, many of the satellite-based products may be spatially and/or temporally inadequate. For example, the MODIS sensors aboard the Terra and Aqua satellites currently provide one or two (two to four with nighttime observations) overpasses per day, depending on the product. Many of the currently available products such as land cover are provided only on a 16-day cycle. While this may prove to be adequate temporal data, future temporal improvements would offer better initial parameterization of BASINS-HSPF.

Another challenge in meeting BASINS needs is the incorporation of NASA gridded data into the HSPF model. Many input variables in BASINS-HSPF are based on the nearest weather station data. In order to ingest NASA precipitation data, for example, a single hourly measurement of precipitation will have to be determined for the entire watershed based on the overlapping and surrounding grid cell values. The grid cell value corresponding to the latitude/longitude of the nearest rain gauge could be used, or it may be more appropriate to average, or “lump” several grid values corresponding to the location of the entire watershed. There are other ways to deal with this issue. HSPF has the capability to automatically subdivide the watershed into sub-basins that can approximate the boundaries of the gridded data. This option needs to be rigorously tested.

Continuity of NASA data products represents another challenge in meeting the needs of BASINS and HSPF. Currently, GPM is a planned precipitation mission that show the great

promise in remotely-sensed components of the water cycle. In the meantime, the NASA's LIS model can provide spatially and temporally accurate data using ground, airborne, and space-based data and short-term modeled results. LIS assimilates data from a variety of sources, including MODIS, NOAA Stage 2 Doppler radar, and weather station data. The LSMs in LIS can also provide short-term predictions of the water cycle. In order to provide useful data for BASINS, it will be necessary to show the decision makers that similar data will continue to be available in the future for use in the DST.

5.0 PARTNERING

5.1 Overview

This project is based on needs documented in the Memorandum of Understanding between NASA and The Environmental Protection Agency for Cooperation in Water, Coastal and Earth Sciences. In this document, NASA agrees to:

1. Support EPA science and technology research, development, transfer, utilization, and commercial efforts within the Research, Economics and Education Mission Area as agreed upon by providing technical expertise for performance, planning, review, or consultation in areas of mutual interest, subject to program priorities and budget constraints.
2. Assist EPA through collaborations to evaluate, verify, validate, and benchmark practical uses of NASA-sponsored observations from remote sensing systems and predictions from scientific research and modeling through the NASA Earth Sciences Enterprise (ESE).

NASA and EPA have identified ten areas of shared goals for improving decision making, policy, and management through beneficial and appropriate use of Earth science data and modeling. Of these ten areas, at least eight are natural extensions of ongoing research and capabilities within the Hydrological Sciences Branch at GSFC. Further collaborations for this project are currently being developed with groups at SSC and several universities that have demonstrated expertise in one or more of these areas.

The combined NASA and EPA teams have identified the highest priority area for possible improvement through the use of NASA Earth science technology as being related to nonpoint source pollution. Details of this collaboration are in the NASA approved project plan, "Water Management Plan: Nonpoint Source Pollution, 2004".

(http://aiwg.gsfc.nasa.gov/esappdocs/progplans/water_ver1-1.pdf).

5.2 Developing Partnerships

The EPA's Office of Water and NASA's Earth Science Enterprise (ESE) entered into a Memorandum of Understanding (MOU) in 2003 to study the use of NASA remote sensing and modeling information to support EPA's water-related programs. Within this framework, NASA/GSFC and EPA developed a project plan (NASA, 2004) under NASA Water Management to study the use of NASA data to improve EPA's water quality program. The

unique capabilities provided by NASA satellite remote sensing and modeling have significant potential to address critical deficiencies for EPA modeling of spatially and temporally variable nonpoint source pollution. This project attempts to leverage the large investment in ESE data to a federal agency with national applications that may provide a significant return for policy making on water quality affecting people's every day lives. A recent Gallup News Service Poll (Saad, 2002) reported the top three out of ten environmental concerns of Americans involve water quality.

As a result of the MOU between NASA HQ, NASA/GSFC along with the EPA Office of Water prepared a five year project plan (NASA, 2004), "BASINS: Nonpoint Source Water Quality" including work with the University of Maryland Center for Environmental Studies and Hunter College that was approved by NASA HQ. NASA/GSFC received funding in 2004 to work with the EPA to further develop relationships and start work. Thus far, the time invested by NASA/GSFC with the EPA has been on study site selections, training, model calibration, and preliminary evaluation studies described in this document.

University of Maryland Center for Environmental Science (UMCES) (PI's P. Townsend and K. Eshleman) UMCES has worked with NASA to test the applicability of new NASA data products to the BASINS watershed assessment tool. The primary emphasis was the study of multispatial scales between capabilities of higher resolution systems (Landsat and IKONOS) and moderate resolution systems such as MODIS for the Chesapeake Bay study sites. The remote sensing expertise of Townsend and water quality/ecosystem expertise of Eshleman made an ideal group to study NASA products to improve BASINS performance.

University of Wisconsin – Dr. Phil Townsend will lead the U. of WISC team. Dr. Townsend focuses on testing and implementing NASA satellite derived land use / land cover and vegetation index data and related phenological changes into HSPF. Dr. Townsend will coordinate with CBP and Dr. Gutiérrez-Magness to complete NASA Evaluation, V&V and Benchmarking reports.

Hunter College, CUNY. (PI Wenge Ni- Meister) Dr. Ni-Meister has significant Land Data Assimilation System (LDAS), data assimilation, and remote sensing expertise to study effects of NASA MODIS and LDAS products on BASINS. This coupled with their strong department work on GIS should enable a thorough analysis of test watersheds using LDAS and satellite data such as from MODIS.

Chesapeake Bay Program, Annapolis Maryland, Gary Shenk (Chesapeake Bay Program (CBP) Office) and Angelica L. Gutiérrez-Magness (UMCP/USGS). EPA and CBP will help coordinate the selection of test sites and watersheds. They will coordinate and provide assistance with setting up HSPF and performing calibrations. Shenk and Gutiérrez-Magness will provide the CBP phase 5 version of HSPF code and sample datasets for the applications team. They will work closely with team in all phases and will participate in informal meetings and quarterly reporting periods.

NASA/GSFC – David Toll (NASA/GSFC) is the Team Leader with assistance from Edwin Engman. GSFC is responsible for coordination of activities between groups. GSFC is also

the lead group evaluating NASA LIS precipitation to BASINS. Joe Nigro provides GIS expertise and assists with BASINS-HSPF runs.

6.0 CONCLUSIONS AND RECOMMENDATIONS

6.1 Findings

The evaluation of potential improvements to the operation of the BASINS streamflow simulation model, HSPF, has pointed to the potential for encouraging and positive improvements. Because HSPF is a lumped parameter model that is highly dependent upon input data, initial conditions and watershed descriptive parameters, there appear to be several opportunities for NASA satellite and modeling products to make significant improvements in HSPF simulations.

6.2 Recommendations

The following is a step by step recommendation for testing and benchmarking our ability to use NASA data and science products to improve the accuracy of the EPA BASINS-based modeling of watershed drainage.

1. Select watersheds to be studied according to the following criteria:

- Period of record for rainfall, stream flow and water quality measurements should be 3 to 5 years.
- No existing reservoirs or lakes above monitoring points
- Watersheds should have been modeled by EPA or a state agency (or contractor) using BASINS and HSPF
- HSPF model parameters are available for use.

2. Select periods of continuous data according to the following criteria:

- Data from both a wet and dry year (based on average rainfall)
- Winter, including snow, if applicable
- Summer, with convective, high rainfall intensity storms
- At least two months of continuous data with significant storms (2-year or greater events)

3. Conduct an HSPF default run using the original forcing input and land and soil input (**Verification**) The **verification** step will involve comparisons of the HSPF model output with the measured flow and this will establish the **baseline benchmark**. The comparisons will involve graphical plots of annual and storm hydrographs and statistical measures of the differences between the models produced and measured streamflow.

4. We will then **validate** our procedures by running similar experiments on the other chosen watersheds within the Chesapeake Bay watershed. The final **verification and validation** step will be making the enhanced version of HSPF available for demonstration using operational NASA data and data products

5. . The last step will be the **benchmarking** and documentation of the performance of the enhanced DSS (in this case, HSPF) by assessing its performance and comparing results with the **baseline**. Measures of efficiency in matching the measured flows as indicated above will be calculated. Statistics such as storm and monthly volumes, time to peak, Nash-Sutcliffe statistic, etc will be used to develop the **BENCHMARK**, against which all future model runs will be compared.

6. Since model calibration may indirectly affect comparisons using NASA data versus traditional comparisons, intermediate comparisons and validations will be conducted. For example use of LIS precipitation will be compared directly to gauge data and surface observations of precipitation in addition to flow related evaluations.

We plan to incrementally drive HSPF with input data from MODIS and LIS to see if we can improve the fit between measured and model derived results. We plan to do similar experiments with improved parameters and improved initial conditions. After seeing which forcings, parameters and initial conditions improve the HSPF/BASINS results, we plan to experiment with combinations.

6.3 Next steps

Our initial study site will be the northeastern Anacostia River and the Patuxent River watersheds in the Chesapeake Bay watershed. Both rivers are in heavily impacted basins, surrounded by urban development of the Baltimore, MD and Washington, D.C. metropolitan areas. The Anacostia flows generally south, draining into the Potomac River, just upstream from the Chesapeake Bay, while the Patuxent flows directly into the Chesapeake Bay.

The Patuxent watershed has been the focus of several recent analyses of nutrient loadings and water quality trends (Boynton et al., 1995; Preston and Summers 1997; Jordan et al., 1999; Weller et al., 2003; and D'Elia et al., 2003) that have addressed the goal of reducing nitrogen and phosphorus loadings to the Chesapeake Bay (including the Patuxent estuary) by 40% by the year 2000. Explosive population growth and rapidly changing land use in the basin continue to make the improvement of water quality in the Patuxent River and estuary an ever-challenging goal; the application of a state-of-the-art water quality model, HSPF, to the basin should contribute in a positive way in achieving these water quality management objectives.

In addition to the Patuxent, we have chosen several other watersheds within the greater Chesapeake Bay basin for our analysis. The watersheds selected include the Mahantango Creek (PA), Little River (VA), Deer Creek (PA), Calfpasture Creek (VA), and the Pocomoke River (DE). These basins represent a sampling of differing land use and topography. Additional basins may be added in the future if we need additional samples to explain and demonstrate results. We are comparing model results with measured data (**Verification**) for the Patuxent watershed (FY2005). This is being done using model calibration parameters suggested by our EPA partners. These calibration parameters represent the **baseline** situation for which future comparisons will be made. Comparisons will be based on visual comparisons against the **baseline** and a series of statistics developed from the HSPF modeling.

The US EPA will coordinate the calibration of HSPF for the test watersheds proposed in this study. Parameterization of HSPF has been accomplished using the PEST tool now included in BASINS. PEST tunes key parameter values to optimize model performance. Because PEST calibration has been undertaken using existing input and driver data sets, the existing calibration may not be optimal for model performance using the new inputs we propose. Alternatively, re-calibration using new data sets will limit comparability of model runs using different input data because all other factors in the model (i.e., parameter values) will not be kept constant. Nevertheless, optimal HSPF performance using the new inputs will require re-calibration. To address this, our research will focus primarily on using the new NASA inputs with the existing calibrated versions of HSPF. However we will also re-calibrate HSPF using PEST and our hypothesized improved inputs. We will then compare HSPF runs using both the new inputs and the existing inputs in the re-calibrated mode. This necessitates comparisons of four sets of HSPF runs:

- (1) original inputs into original calibration HSPF
- (2) new NASA inputs into original calibration HSPF
- (3) new NASA inputs into HSPF recalibrated with new inputs
- (4) original inputs into HSPF recalibrated with new inputs

7.0 REFERENCES

- Anderson, J.R., E. Hardy, J. Roach, and R. Witmer. 1976. A land use and land cover classification system for use with remote sensing data. *U.S. Geological Survey Professional Paper 964*, Washington, DC.
- Bicknell, B.R., J.C. Imhoff, J.L. Kittle, A.S. Donigan, and R.C. Johanson. 1997. Hydrological simulation program FORTRAN (HSPF): user's manual for release 11. *EPA-600/R-97-080*, U.S. Environmental Protection Agency, Athens, GA.
- Boynton, W.R., J.H. Garber, R. Summers, and W.M. Kemp. 1995. Inputs, transformations, and transport of nitrogen and phosphorous in Chesapeake Bay and selected tributaries. *Estuaries* 18: 285-314.
- Civco, D.L., J.D. Hurd, E.H. Wilson, C.L. Arnold, and M.P. Prisloe. 2002. Quantifying and describing urbanizing landscapes in the Northeast United States. *Photogrammetric Engineering and RemoteSensing* 68: 1083-1090.
- Cosgrove, B.A., D. Lohmann, K.E. Mitchell, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, C. Marshall, J. Sheffield, Q. Duan, L. Luo, R.W. Higgins, R.T. Pinker, J.D. Tarpley, and J. Meng. 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research* 108 (D22): Art. No. 8842.
- D'Elia, C.F., W.R. Boynton, and J.G. Sanders. 2003. A watershed perspective on nutrient enrichment, science, and policy in the Patuxent River, Maryland: 1960-2000. *Estuaries* 26: 171-185.
- Donigan, A.S. Jr., B.R. Bicknell, and J.C. Imhoff. 1995. Hydrological Simulation Program – FORTRAN (HSPF). In *Computer Models of Watershed Hydrology* (V.P. Singh, ed.), Chapter 12, pp. 395-442, Water Resources Publications, Littleton, CO.
- Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler. 2002. The National Elevation Dataset. *Photogrammetric Engineering and Remote Sensing* 68: 5.

- Hansen, M.C., R.S. DeFries, J.R.G. Townshend, R. Sohlberg, C. Dimiceli, and M. Carroll. 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment* 83: 303-319.
- Hargrove, W.W., and F.M. Hoffman. 2004. The potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management*, 34(5): 39-60.
- Hargrove, W.W., and R.J. Luxmoore. 1998. A Clustering Technique for the Generation of Customizable Ecoregions. *Proceedings, ESRI Arc/INFO Users Conference*, Palm Springs, CA.
- Jantz, C.A., S.J. Goetz, and M.A. Shelley. 2004. Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan region. *Environment and Planning B: Planning and Design* 31: 251-271.
- Jordan, T.E., D.E. Weller, and D.L. Correll. 1999. Sources of nutrient inputs to the Patuxent River estuary. *Estuaries* 26:226-243.
- Mitchell, K.E., D. Lohmann, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, B.A. Cosgrove, J. Sheffield, Q.Y. Duan, L.F. Luo, R.W. Higgins, R.T. Pinker, J.D. Tarpley, D.P. Lettenmaier, C.H. Marshall, J.K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B.H. Ramsay, and A.A. Bailey. 2004. The multi-institution North America Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research* 109 (D7): No. D07S90.
- NASA, 2004. "BASINS: Nonpoint Source Water Quality" Project Plan approved by Water Management Applications Program, NASA-HQ, Washington, DC.
(http://aiwg.gsfc.nasa.gov/esappdocs/projplans/BASINS_WMProject.doc)
- Preston, S.D., and R.M. Summers. 1997. Estimation of nutrient and suspended-sediment loads in the Patuxent River basin, Maryland, water years 1986-1990. *Water Resources Investigations Report 96-4175*, U.S. Geological Survey, Baltimore, MD, 69 pp.
- Saad, L. 2002. "Americans Sharply Divided on Seriousness of Global Warming." Gallup News Service, Princeton, NJ.
- Smith, B., and D. Sandwell. 2003. Accuracy and resolution of shuttle radar topography mission data. *Geophysical Research Letters* 30, No. 9.
- Smith, R.A., R.B. Alexander, and M.G. Wolman. 1987. Water-Quality Trends in the Nations Rivers. *Science* 235: 1607-1615.
- USDA, 1991. State Soil Geographic (STATSGO) Data Base. Natural Resources Conservation Service, national Soil Survey Center, Pub. No. 1492.
- USEPA. 2001. Better Assessment Science Integrating point and Nonpoint Sources BASINS Version 3.0 User's Manual, EPA-823-B-01-001, U.S. Environmental Protection Agency, Office of Water, Washington, DC.
- U.S. Governmental Accountability Office (USGAO), (2000), "Water Quality, Key EPA and State Decisions limited by Inconsistent and Incomplete Data", Report to the Chairman, Subcommittee on Water Resources and Environment, Committee on Transpiration and Infrastructure, House of Representatives, March 2000.

- Vogelmann, J.E., S.M. Howard, L.M. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel. 2001. Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing* 67: 650-662.
- Wang, Y.Q., and X.S. Zhang. 2004. A SPLIT model for extraction of subpixel impervious surface information. *Photogrammetric Engineering and Remote Sensing* 70: 821-828.
- Weller, D.E., T.E. Jordan, and D.L. Correll. 2003. Effects of land-use change on nutrient discharges from the Patuxent River watershed. *Estuaries* 26: 244-266.
- Zhang, X.Y., M.A. Friedl, C.B. Schaaf, A.H. Strahler, J.C.F. Hodges, F. Gao, B.C. Reed, and A. Huete. 2003. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment* 84: 471-475.

8.0 ABBREVIATIONS AND ACRONYMS

ALI	Advanced Land Imager
AMSR-E	Advanced Microwave Scanning Radiometer-EOS
ArcView	ESRI GUI-based GIS
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
BASINS	Better Assessment Science Integrating Point and Nonpoint Sources
BMP	Best Management Practices
BRDF	Bidirectional Reflectance Distribution Function
CBP	Chesapeake Bay Program
CUNY	City University of New York
DEM	Digital Elevation Model
DO	Dissolved Oxygen
DSS	Decision Support System
DST	Decision Support Tool
EO-1	NASA's Earth Observing-1 satellite
EOS	NASA's Earth Observing System
EPA	U.S. Environmental Protection Agency
ESE	NASA's Earth Science Enterprise
ESRI	Environmental Systems Research Institute, Inc.
ET	Evapotranspiration
FPAR	Fraction of Photosynthetically Active Radiation
FSL	NOAA's Forecast Systems Laboratory
GAPP	NOAA's GEWEX Americas Prediction Project
GEWEX	Global Energy and Water Cycle Experiment
GIS	Geographic Information System
GMAO	NASA's Global Modeling and Assimilation Office
GPM	NASA's Global Precipitation Measurement Mission
GSFC	NASA's Goddard Space Flight Center
HQ	NASA Headquarters
HSPF	Hydrological Simulation Program - FORTRAN
IGBP	International Geosphere-Biosphere Programme
LAI	Leaf Area Index
LDAS	Land Data Assimilation System

LIS	Land Information System
LSM	Land Surface Model
MODIS	Moderate Resolution Imaging Spectroradiometer
MOU	Memorandum of Understanding
MVI	Modified Vegetation Index
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
NEXRAD	NOAA's Next Generation Doppler Radar
NLDAS	North American Land Data Assimilation System
NOAA	National Oceanic and Atmospheric Administration
PEST	Parameter Estimation and tool for model calibration
PI	Principal Investigator
PLOAD	BASINS Pollutant Loading Application
Pot. ET	Potential Evapotranspiration
QUAL2E	Enhanced Stream Water Quality Model
RUC	Rapid Update Cycle model
SRTM	Shuttle Radar Topography Mission
SSC	NASA's Stennis Space Center
STATSGO	State Soil Geographic Database
SWAT	Soil and Water Assessment Tool
SWIR	Short Wavelength Infrared
SWM	Stanford Watershed Model
TIR	Thermal Infrared
TM	Landsat Thematic Mapper
TMDL	Total Maximum Daily Load
TRMM	Tropical Rainfall Measuring Mission
U Wind	East-West component of wind vector
UMCES	University of Maryland Center for Environmental Science
UMCP	University of Maryland, College Park
UMD	The University of Maryland
USGAO	U.S. Government Accountability Office
USGS	U.S. Geological Survey
V Wind	North-South component of wind vector
V&V	Verification and Validation
VNIR	Visible and Near Infrared

Appendix A. The NASA Land Information System (LIS)

The Land Information System infrastructure (LIS; Peters-Lidard et al., 2004; Kumar et al., 2006) unifies and extends the capabilities of the $\frac{1}{4}$ degree Global Land Data Assimilation System (Rodell et al. 2004) and the $\frac{1}{8}$ degree North American LDAS (NLDAS; Mitchell et al., 2004) in a common software framework capable of ensemble land surface modeling on points, regions or the globe at spatial resolutions from 2.5 degrees down to 1km and finer. LIS is led by the Hydrological Sciences Branch at NASA Goddard Space Flight Center. The 1km and finer resolution capability of LIS allows it to take advantage of the latest EOS-era observations, such as MODIS leaf area index, surface a, snow cover, albedo and surface temperature and AMSR-E snow water equivalent. LIS builds upon the capabilities of the $\frac{1}{4}$ -degree Global Land Data Assimilation System (GLDAS; Rodell et al. 2004; <http://ldas.gsfc.nasa.gov>) and the $\frac{1}{8}$ -degree North American Land Data Assimilation System (NLDAS; Mitchell et al. 2004; <http://ldas.gsfc.nasa.gov>) to determine energy and water states (e.g. snow depth and soil moisture) and fluxes (e.g. evaporation, transpiration and runoff) at 1-km and finer spatial resolutions, and at one-hour and finer temporal resolutions. The 1-km capability of LIS allows it to take advantage of the latest EOS-era observations, such as MODIS leaf area index, snow cover and surface temperature, at their full resolution. Figure A-1 illustrates the data integration and data assimilation capabilities for water resources applications.

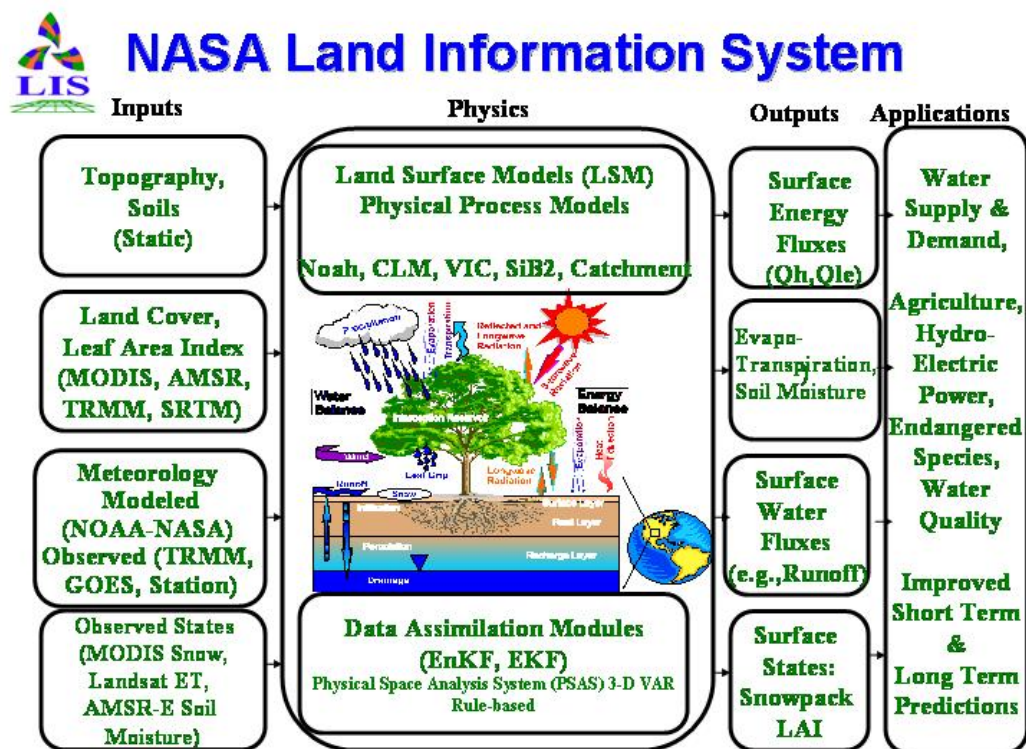


Figure A-1. The NASA-GSFC Land Information System integrating a range of spatial remote sensing and ancillary products in to a suite of land surface models with the capability for data assimilation serving a wide range of water resources applications.

LIS features a high-performance and flexible design, provides infrastructure for data integration and assimilation, and operates primarily on an ensemble of land surface models for execution over user-specified regional or global domains. The LIS framework is designed using advanced software engineering principles to enable the reuse and community sharing of modeling tools, data resources, and assimilation algorithms. LIS provides generic, model-independent support for high performance computing, resource management, data handling, inter-language support and other functions. The LIS software is designed within an object-oriented framework, with explicit abstract interfaces defined for customization and extension to different applications. These extensible functionalities, or “plugging,” in LIS include interfaces for the incorporation of new domains, land surface models, land surface parameters, meteorological input schemes and data assimilation algorithms. As the extensible components are designed to remain independent from specific models and algorithms, the component-style specification of the system allows rapid prototyping and development of applications.

The use of observation-driven land models and data assimilation is a fundamental principle of LIS and enables the communication of these products to DST models and solutions. A huge volume of land surface observations are or may be operationally sensed from space, including surface temperatures, vegetation conditions, snow states, albedo, longwave and solar radiation, precipitation, surface moisture, freeze/thaw state, runoff, total water storage, and elevation, among others.

Advances in the understanding of soil-water dynamics, plant physiology, small-scale meteorology, and the hydrology that control biosphere–atmosphere interactions have spurred the development of Land Surface Models (LSMs), whose aim is to represent properly the transfer of mass, energy, and momentum between a vegetated surface and the atmosphere. LSM predictions are regular in time and space, but these predictions are influenced by errors in model structure, input variables, parameters, and inadequate treatment of sub-grid scale spatial variability. Consequently, LSM predictions are significantly improved through observation constraints. Our team has adopted an “ensemble physics” land surface modeling philosophy to enable straightforward collaboration with operational weather, climate, and decision support partners. The land surface models currently incorporated in LIS include the Variable Infiltration Capacity (VIC) model (Lang et al. 1996), version 2 of the Common Land Model (CLM2; Dai et al. 2003), the Community Noah LSM (Eke et al. 2003), Mosaic (Kosher and Suarez 1996), and the Hyssop model (Sud and Mocko 1999; Mocko et al. 1999). Figure A-2 below shows a diagram of the land surface modeling concept.

LIS also provides a web-based user interface that accesses data mining, numerical modeling, and visualization tools. The LIS is packaged into a portable system for download together with a small database to an independent system for use in smaller applications, or used on a centralized server for large applications. And LIS defines land surface modeling and assimilation standards, enabling straightforward coupling to Earth System Modeling Frameworks (ESMF).

Land Surface Modeling Concept

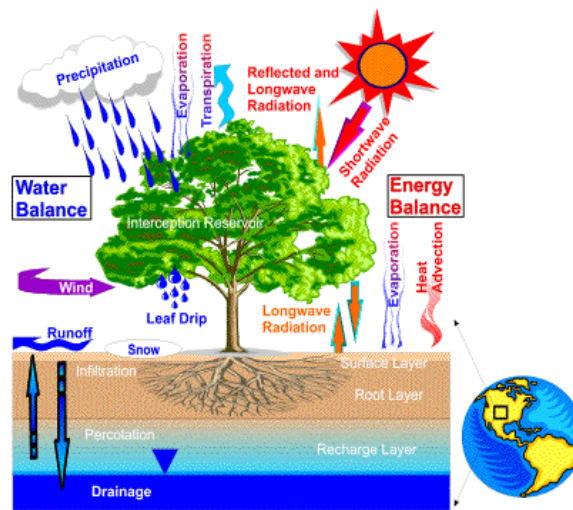


Figure A-2. Illustrates the complex land-atmosphere inter-relationships of soil-water dynamics, plant physiology, small-scale meteorology, and the hydrology that control biosphere-atmosphere interactions.

The main software components of LIS are:

- LIS driver: A model control and input/output system that executes multiple offline land surface models over regional or global grids/tiles at spatial resolutions down to 1km.
- Land surface models: The LIS source code currently includes 3 different land surface models, namely, The NCAR Community Land Model (CLM); The community Noah land surface model (Noah) ; and The Variable Infiltration Capacity model (VIC) .

The data used by LIS include:

- Parameter data : Properties of the land surface that change on time-steps of a day or longer, e.g., soil, land cover, topography.
- Forcing data : Atmospheric inputs to the land surface models, including precipitation, radiation, and surface winds, temperature, pressure and humidity.

In order to predict water, energy and biogeochemical processes using (typically 1-D vertical) partial differential equations, land surface models require three types of inputs: 1) Initial conditions, which describe the initial state of the land surface; 2) Boundary conditions, which describe both the upper (atmospheric) fluxes or states also known as "forcings" and the lower (soil) fluxes or states; and 3) Parameters, which are a function of soil, vegetation, topography, etc., and are used to solve the governing equations.

Output from the land surface models translates to variables in “ALMA” standard format such as soil moisture, surface runoffs, and canopy conductance. The "Get LIS Data link” at the LIS website (<http://lis.gsfc.nasa.gov>) allows registered users to download data via HTTP, or visualize the data using the LIS Land Explorer (LE). The LIS web site also has a HYPERLINK "<http://lis.gsfc.nasa.gov/Gallery/index.shtml>" Gallery section that will contain plots of typical model runs. For more information on the LIS design, please review the Software Design Document on the LIS web page (<http://lis.gsfc.nasa.gov>).

The goal of the user interface components in LIS is to allow the interactive, flexible use of the LIS products to the end users. The visualization capabilities in LIS are built on a multi-tiered client-server system architecture. LIS is designed to encourage the reuse and community sharing of scientific land-atmosphere modeling algorithms. The interoperable features in LIS also include the reuse and participation with other Earth system modeling groups. The LIS system software architecture of LIS is depicted in Figure A-3. The top layer handles operations related to the overall program control and a number of generic tools.

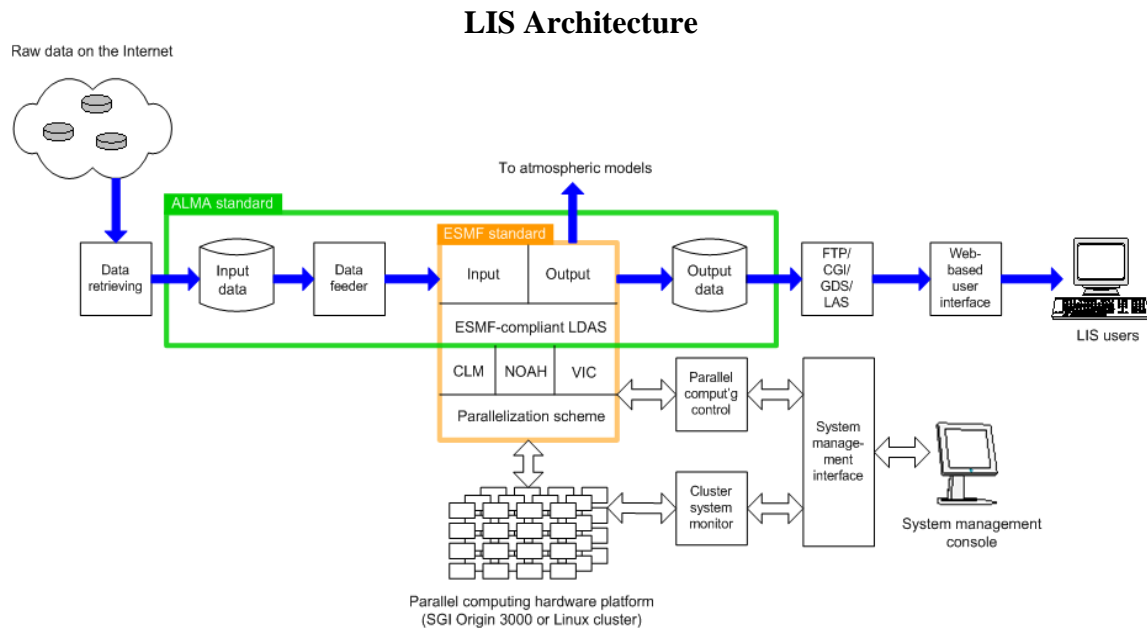


Figure A-3. Components of the LIS system software architecture.

A-1. References

Dai, Y., X. Zeng, R.E. Dickinson, I. Baker, G.B. Bonan, M.G. Bosilovich, A.S. Denning, P.A., Dirmeyer, P.R. Houser, G.N. Keith, W. Oleson, C.A., Schlosser and Z-L Yang, 2003: The Common Land Model. *Bull. Amer. Meteor. Soc.*, 84, pp. 1013-1023.

Koster, R. D., and M. J. Suarez, 1996: Energy and water balance calculations in the Mosaic LSM. NASA Technical Memorandum 104606, no. 9, 76 pp.

- Kumar, S. V., C. D. Peters-Lidard, Y. Tian, J. Geiger, P. R. Houser, S. Olden, L. Lighty, J. L. Eastman, P. Dirmeyer, B. Doty, J. Adams, E. Wood and J. Sheffield, 2006: LIS—An interoperable framework for high resolution land surface modeling. *Environmental Modeling and Software*, in press, (available on-line at <http://lis.gsfc.nasa.gov>).
- Liang, X., D. Lettenmaier, and E. Wood, 1996. One-dimensional statistical dynamic representation of subgrid variability of precipitation in the two-layer variable infiltration capacity model, *J. Geophys. Research*, 101(D16):21403-21422.
- Mitchell, K. E., and coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, 109, D07S90, doi: 10.1029/2003JD003823.
- Mocko, D. M., G. K. Walker, and Y. C. Sud, 1999: New snow-physics to complement SSiB. Part II: Effects on soil moisture initialization and simulated surface fluxes, precipitation, and hydrology of GEOS II GCM. *J. Meteor. Soc. Japan*, 77 (1B), 349-366.
- Peters-Lidard, C.D., S.V. Kumar, Y. Tian, J.L. Eastman, and P.R. Houser, 2004: Global urban-scale land-atmosphere modeling with the Land Information System. Presented at the Symposium on Planning, Now-casting, and Forecasting in the Urban Zone during the 84th Annual Meeting of the American Meteorological Society, 10-16 January 2004, Seattle WA.
- Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, *Bull. Amer. Meteor. Soc.*, 85 (3), 381–394, 2004
- Sud, Y.C., and D.M. Mocko, 1999: New snow-physics to complement SSiB. Part I: Design and evaluation with ISLSCP Initiative I datasets. *J. Meteor. Soc. Japan*, v. 77, pp. 335-348.